

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-147119

(43)公開日 平成9年(1997)6月6日

(51)Int.Cl.<sup>6</sup>

G 0 6 T 7/60  
1/00

識別記号

府内整理番号

F I

G 0 6 F 15/70  
15/62

技術表示箇所

3 5 0 H  
3 8 0

審査請求 未請求 請求項の数11 OL (全13頁)

(21)出願番号 特願平7-309083

(22)出願日 平成7年(1995)11月28日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 榎本暢芳

神奈川県川崎市幸区柳町70番地 株式会社

東芝柳町工場内

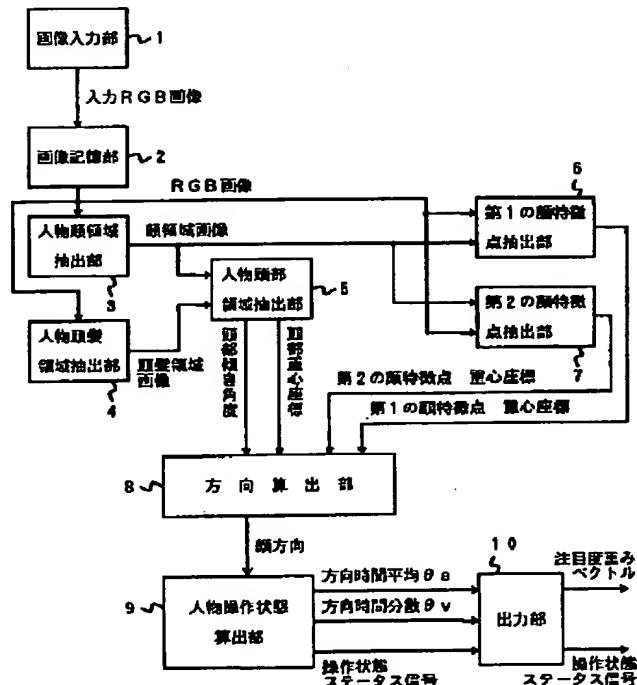
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】人物注目方向検出方式

(57)【要約】

【課題】操作人物の状態を自動的に把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示などを行なうことが可能となる人物注目方向検出方式を提供する。

【解決手段】画像入力部1で操作人物の顔領域を含む画像を入力し、この入力された画像から人物顔領域抽出部3で顔領域を抽出し、人物頭髪領域抽出部4で頭髪領域を抽出し、これら抽出された顔領域と頭髪領域とから人物頭部領域抽出部5で頭部領域を抽出し、人物顔領域抽出部3で抽出された顔領域内において第1の顔特徴点抽出部6で口領域を抽出し、第2の顔特徴点抽出部7で目領域を抽出し、これら抽出された頭部領域と口領域および目領域との位置関係から方向算出部8で顔の方向を求め、この求めた顔方向の変化状況から人物操作状態算出部9で人物の操作状態を推定し、その推定結果を出力部10で外部へ出力する。



## 【特許請求の範囲】

【請求項 1】 機器を操作する人物の顔領域を含む画像を入力し、この入力された画像から操作対象人物の顔領域および顔内部の特徴点位置を特定し、これら特定した顔領域および顔内部の特徴点位置を用いて顔の方向とその時間的な統計を検出することによって、操作人物の注目方向と感情を含む操作状態とを推定することを特徴とする人物注目方向検出方式。

【請求項 2】 機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域内において少なくとも 2 種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段と、を具備したことを特徴とする人物注目方向検出方式。

【請求項 3】 機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域内において少なくとも 2 種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段と、この人物操作状態算出手段で推定された人物の操作状態を示す情報を外部へ出力する出力手段と、を具備したことを特徴とする人物注目方向検出方式。

【請求項 4】 前記カラー画像は R (赤) , G (緑) , B (青) のカラー画像であり、前記人物顔領域抽出手段は、前記画像記憶手段に記憶された R, G, B のカラー画像から顔領域を抽出するために、前記 R, G, B のカラー画像を特定表色系のカラー画像に変換し、頻度分布上で前記顔領域の有する濃度が他の領域の濃度と有意差を持つ特定の 1 種類の色成分画像を使用して、頻度分布上でのクラス分けによる領域候補の抽出を行なうとともに、あらかじめ求められた画面内の人物の顔および頭の全画面に対する面積比率による前記抽出された領域候補の選択を行なうこととする請求項 2 および 3 記載の人物注目方向検出方式。

【請求項 5】 前記カラー画像は R (赤) , G (緑) , B (青) のカラー画像であり、前記人物頭髪領域抽出手段は、前記画像記憶手段に記憶された R, G, B のカラー画像から頭髪領域を抽出するために、前記 R, G, B のカラー画像を特定表色系のカラー画像に変換し、頻度分布上で前記頭髪領域の有する濃度が他の領域の濃度と有意差を持つ特定の 1 種類の色成分画像を使用して、頻度分布上でのクラス分けによる領域候補の抽出を行なうとともに、あらかじめ求められた画面内の人物の顔および頭の全画面に対する面積比率による前記抽出された領域候補の選択を行なうこととする請求項 2 および 3 記載の人物注目方向検出方式。

【請求項 6】 前記顔特徴点抽出手段は、抽出する 2 種類の顔特徴点のうちの 1 種類として口領域を用い、その口領域を抽出するために、頻度分布上で口領域の濃度が他の領域の濃度に対して有意差を持つ特定の表色系の 1 種類の色成分画像を前記画像記憶手段に記憶された画像から抽出することにより口候補領域の抽出を行なうとともに、口候補領域の面積、位置、幅とからなるコスト評価関数の最小化により前記抽出された口候補領域の選択を行なうことを特徴とする請求項 2 および 3 記載の人物注目方向検出方式。

【請求項 7】 前記顔特徴点抽出手段は、抽出する 2 種類の顔特徴点のうちの 1 種類として目領域を用い、その目領域を抽出するために、頻度分布上で目のうちの瞳領域の濃度が他の領域の濃度に対して有意差を持つ特定の表色系の 1 種類の色成分画像を前記画像記憶手段に記憶された画像から抽出することにより目候補領域の抽出を行なうとともに、目候補領域外における瞳相当領域の面積と顔との対象軸からの距離と標準的人物の目全体面積に対する瞳領域面積の比率とからなるコスト評価関数の最小化により前記抽出された目候補領域の選択を行なうことを特徴とする請求項 2 および 3 記載の人物注目方向検出方式。

【請求項 8】 前記人物操作状態算出手段は、顔方向を一定の離散時間で求め、これらの時間的な変化パターンや統計量を算出し、あらかじめ複数の被験者を対象として求められた被験者の操作状態を示す標準的な参照データ

タと比較することにより人物の操作状態を推定することを特徴とする請求項2および3記載の人物注目方向検出方式。

**【請求項9】** 前記出力手段は、前記方向算出手段で求められたある瞬間の顔方向に対して、前記人物操作状態算出手段で算出された顔方向の時間変化統計量を用いて人物のその方向への注目度を推定することにより、人物の操作状態を方向ベクトルと注目度という2つの量として外部へ出力することを特徴とする請求項8記載の人物注目方向検出方式。

**【請求項10】** 機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域内において少なくとも2種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段と、この人物操作状態算出手段で推定された人物の操作状態を示す情報をグラフィカルユーザインタフェース装置へ出力する出力手段と、を具備したことを特徴とする人物注目方向検出方式。

**【請求項11】** 前記グラフィカルユーザインタフェース装置は、前記出力手段が表示する人物の操作状態を示す情報を受信することにより、操作人物が操作方法がわからずには迷っている状態では、指示メッセージをより細かく出力するようにするといった指示メッセージの内容の変化や、指示メッセージの表示色を目立つようにするといった指示メッセージの方法の変化を起こさせるといった、インタフェースのダイナミックかつ状態に応じた提供を行なうことを特徴とする請求項10記載の人物注目方向検出方式。

#### 【発明の詳細な説明】

##### 【0001】

**【発明の属する技術分野】** 本発明は、たとえば、現金自動入出金装置や券売機などの会社システム機器、あるいは、パーソナルコンピュータやワードプロセッサなどのオフィスオートメーション機器などにおいて、操作人物の注目している方向として人物の顔の向いている方向を

検出し、注目度や操作人物の状態を推定する人物注目方向検出方式に関する。

##### 【0002】

**【従来の技術】** 従来、この種の技術として以下に説明する2つの方式が知られている。

(1) 目尻および口の端点の4特徴点が等脚台形をなすという仮定により、顔方向を推定した例(文献[1]青山宏、河越正弘：顔の対称性を利用した視線感知法、情処研報CV、VOL 61、NO.2参照)、および、顔を各点間距離が既知の3特徴点からなる三角形と仮定し、各点の3次元位置を求め、顔方向を推定した例(文献[2]大村和典、伴野明、小林幸雄：単眼視画像による顔の向き検出法の指示入力への応用、信学論、VOL J72-D-11 NO.9、PP.1441-1447参照)がある。

**【0003】** (2) 頭全体および顔内部の高輝度変化領域を抽出し、前者と後者との重心位置のずれによった顔方向の検出を行ない、実時間による顔方向識別の検討を行なった例(文献[3]島田聰：顔画像からの注視方向の実時間識別法、信学技法IE91-64、PP.17-23参照)、さらに、後者領域の代わりとして特定色相方向彩度によって検出した目および口領域を使用し、照明条件や顔の凹凸による変動に対するロバスト性の向上を検討した例

(文献[4]島田聰、安達文夫、石井健一郎：特定色相方向の彩度による顔方向の識別、1993年春信学全大、D-514、PP.7-275参照)がある。

##### 【0004】

**【発明が解決しようとする課題】** 上記(1)の両方式によれば、特徴点の位置が高精度で検出されていることを前提としており、一般には、環境変動によって生ずる特徴点位置の精度に対するロバスト性や、特徴点を正確に抽出するために費やされる処理時間の点で問題が多い。

**【0005】** また、上記(2)の方式では、上述のように、照明条件や顔の凹凸による変動に対するロバスト性の向上を検討してはいるが、目に髪がかかる場合や、顔が横方向を向いたときなどの変動によって、特徴点位置が正確に検出されないなどの状況に対する考慮がなされていない。

**【0006】** そこで、本発明は、操作人物の状態を自動的に把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示などを行なうことが可能となる人物注目方向検出方式を提供することを目的とする。

**【0007】** また、本発明は、グラフィカルユーザインタフェース装置において、高度なグラフィカルユーザインタフェースが構築可能となる人物注目方向検出方式を提供することを目的とする。

##### 【0008】

**【課題を解決するための手段】** 本発明の人物注目方向検出方式は、機器を操作する人物の顔領域を含む画像を入力し、この入力された画像から操作対象人物の顔領域および顔内部の特徴点位置を特定し、これら特定した顔領

域および顔内部の特徴点位置を用いて顔の方向とその時間的な統計を検出することによって、操作人物の注目方向と感情を含む操作状態とを推定することを特徴とする。

【0009】また、本発明の人物注目方向検出方式は、機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域内において少なくとも2種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段とを具備している。

【0010】また、本発明の人物注目方向検出方式は、機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域内において少なくとも2種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段とを具備している。

【0011】さらに、本発明の人物注目方向検出方式は、機器を操作する人物の顔領域を含むカラー画像を連続的に入力する画像入力手段と、この画像入力手段で入力された画像を一時記憶する画像記憶手段と、この画像記憶手段に記憶された画像から人物の顔の領域を抽出する人物顔領域抽出手段と、前記画像記憶手段に記憶された画像から人物の頭髪の領域を抽出する人物頭髪領域抽出手段と、前記人物顔領域抽出手段で抽出された顔領域と前記人物頭髪領域抽出手段で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出手段と、

前記人物顔領域抽出手段で抽出された顔領域内において少なくとも2種類の顔特徴点位置を抽出する顔特徴点抽出手段と、前記人物頭部領域抽出手段で抽出された頭部領域と前記顔特徴点抽出手段で抽出された顔特徴点との位置関係から顔の方向を求める方向算出手段と、この方向算出手段で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出手段と、この人物操作状態算出手段で推定された人物の操作状態を示す情報をグラフィカルユーザインタフェース装置へ出力する出力手段とを具備している。

【0012】本発明によれば、入力された操作人物の顔領域を含む画像から操作対象人物の顔領域および顔内部の特徴点位置を特定し、それらを用いて顔の方向とその時間的な統計を検出することによって操作人物の注目方向と感情を含む操作状態とを推定することにより、操作人物の状態を自動的に把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示などを行なうことが可能となる。

【0013】すなわち、たとえば、現金自動入出金装置や券売機などの会社システム機器、あるいは、パソコン用コンピュータやワードプロセッサなどのオフィスオートメーション機器などにおいて、一般ユーザが操作方法を理解できないための問題があるが、本発明方式を用いることにより、操作人物の状態を機器側が把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示（ヘルプメニュー）などを行なうことが可能となる。

【0014】また、現在、パソコン用コンピュータやワークステーションなどのグラフィカルユーザインタフェース装置は、グラフィカルユーザインタフェース（GUI）を備えていて、マウス操作を主体とした操作ができるが、本発明方式を用いることにより、注目領域による操作ボタンの処理や、注目領域の表示を変更（拡大、表示色の変更など）することによる、高度なグラフィカルユーザインタフェースが構築可能となる。

【0015】

【発明の実施の形態】以下、本発明の実施の形態について図面を参照して説明する。なお、本実施の形態では、ビデオカメラから入力された画像中に含まれる操作人物の注目している方向として操作人物の顔の向いている方向を検出し、注目度や操作人物の状態（操作方法に迷っている、操作に疲れてきた等）を推定する装置に適用した場合について説明する。

【0016】図1は、本実施の形態に係る人物注目方向検出方式が適用される装置の構成を示している。この装置は、カラーのビデオカメラと、その出力をデジタル化するA/D変換器とからなり、機器（たとえば、パソコン用コンピュータ）を操作する人物の顔領域を含むカラー画像を時間連続的に入力する画像入力部1、画像入力部1により数フレームごとに入力されたR（赤）、G（緑）、B（青）各8ビットのデジタル画像データを一

時記憶する画像記憶部2、画像記憶部2に記憶されたカラー画像データ（入力画像）から人物の顔の領域を抽出する人物顔領域抽出部3、画像記憶部2に記憶されたカラー画像データから人物の頭髪の領域を抽出する人物頭髪領域抽出部4、人物顔領域抽出部3で抽出された顔領域と人物頭髪領域抽出部4で抽出された頭髪領域とから人物の頭部領域を抽出する人物頭部領域抽出部5、人物顔領域抽出部3で抽出された顔領域において顔特徴点として口領域を抽出する第1の顔特徴点抽出部6、人物顔領域抽出部3で抽出された顔領域において顔特徴点として目領域を抽出する第2の顔特徴点抽出部7、人物頭部領域抽出部5で抽出された頭部領域と第1、第2の顔特徴点抽出部6、7で抽出された2種類の顔特徴点との位置関係から顔の方向を求める方向算出部8、方向算出部8で求められた顔方向の変化状況から人物の操作状態を推定する人物操作状態算出部9、および、人物操作状態算出部9で推定された人物の操作状態を示す情報をグラフィカルユーザインタフェース装置としてのパソコン 컴퓨터などへ出力する出力部10によって構成されている。

【0017】すなわち、本実施の形態の装置は、画像入力部1により数フレームごとに入力されたRGB各8ビットのデジタル画像データは、画像記憶部2に一時格納される。画像記憶部2に格納された入力画像に対して、後で詳細を説明するように、人物顔領域抽出部3、人物頭髪領域抽出部4、および、人物頭部領域抽出部5をそれぞれ適用することにより、人物の顔領域画像、頭部領域画像の重心位置、および、投影平面内での頭部の傾き角をそれぞれ抽出する。

【0018】次に、こうして求められた顔領域画像内部について、第1の顔特徴点抽出部6、第2の顔特徴点抽出部7を用いることで、本来顔の正面に位置すると考えられる顔内部の器官（ここでは、第1の特徴点として

$$\begin{aligned} Y &= 0.30 \cdot R + 0.59 \cdot G + 0.11 \cdot B \\ I &= 0.60 \cdot R - 0.28 \cdot G - 0.32 \cdot B \\ Q &= 0.21 \cdot R - 0.52 \cdot G + 0.31 \cdot B \dots \quad (1) \end{aligned}$$

ただし、R、G、B、Y、I、Qは、それぞれカラー画像のR、G、B、Y、I、Q成分を示す。

【0024】次に、上述の処理によって抽出された1成分画像のみの濃度ヒストグラムをとる。ここでは、横軸を濃淡値（0～255）、縦軸を頻度としたものとして濃度ヒストグラムを作成する。

【0025】さらに、この作成した濃度ヒストグラムを濃度軸上で3領域（クラス）に領域分割し、濃度が中間のクラスを選択して肌色領域を抽出する。このときのクラス分割手法としては、K-平均アルゴリズムなどの一般に知られているクラスタリングのアルゴリズムを使用して差し支えない。

【0026】以上の処理によって、入力画像中から肌色をした画像領域を抽出することが可能であるが、入力

口、第2の特徴点として目を使用）の投影面中での位置を検出するとともに、各特徴点の重心位置をそれぞれ計算する。

【0019】この後、方向算出部8によって、上記のように算出された頭部の重心位置、平面内での傾き角、および、特徴点の重心位置から顔方向を計算し、さらに、人物操作状態算出部9によって、各時刻で求められる顔の方向の時間的統計分布を求めて、人物の注目の度合い、および、操作人物の状態（操作方法に迷っている、操作に疲れてきた等）を推定する。

【0020】この推定結果は、一定時間ごとに出力部10によって、算出された顔方向のベクトルに対して注目度に応じた重みを伴わせたデータと、操作状態のデータ（注目している、操作に迷っている、操作を中断している）として、パソコンコンピュータなどへ出力される。

【0021】以下、各部の処理について詳細に説明する。まず、人物顔領域抽出部3の処理について、図2に示すフローチャートを参照して説明する。一般に、YIQ表色系（文献[5] 高木、下田：画像解析ハンドブック PP.103 参照）における、1成分画像の濃度頻度分布（濃度ヒストグラム）には、肌色部分とそれ以外の部分とで形状に顕著な差があることが知られている。すなわち、濃度ヒストグラムの濃淡値軸の中位の部位で、他部位との形状の異なる領域に人物の肌色に近い色が配置される。

【0022】本実施の形態では、この知識を用い、まず、画像記憶部2内のRGBのカラー画像をYIQ表色系のカラー画像に変換する。このときの変換式は、一般に用いられている以下のようない式（1）で差し支えない。

【0023】

像の中には、たとえば、手などのように、顔以外で肌色をした領域が存在する可能性があるので、ここでは、あらかじめ全画面に対する顔領域の面積比率を求めておき、上記で求まった肌色の領域のうちで、面積比率が顔のそれに最も近いものを選択することで、顔領域を抽出する。ここで肌色領域の面積の計算は、連結領域のラベリング手法（文献[6] 高木、下田：画像解析ハンドブック PP.578 参照）や射影抽出の閾値処理による手法などを使用して差し支えない。

【0027】上記処理によって抽出された顔画像領域は、顔部分の濃淡値を「1」、顔以外の部分の濃淡値を「0」とした2値デジタル画像（顔領域画像）として出力される。

【0028】次に、人物頭髪領域抽出部4の処理につい

て、図3に示すフローチャートを参照して説明する。頭髪領域の抽出には、画面内での人物の頭髪領域の画面全体に対する面積比率を既知として利用し、「黒領域で、画面全体に対する面積比率が既知の値に最も近く、顔領域の上部に位置するもの」を検出する。

【0029】この検出においても、顔と同様に、ある種の色差画像において、ヒストグラム形状に特徴を求ることとし、Y成分画像を数クラスに分割したときの最小濃度のクラスを抽出して使用する。これは、変換が容易な色差成分であるY, I, Qのうちで、Y成分での頭髪部分画素に対する濃度頻度分布（ヒストグラム）が最も特徴的であり、クラス分離しやすいためである。

【0030】したがって、ここでは、上述の人物顔領域抽出部3の処理と同様に、まず、入力画像に対してRG B→YIQ変換を行なう。次に、上述の処理にて抽出されたY成分画像のみの濃度ヒストグラムをとり、これを濃度軸上で領域分割し、濃度が最小の領域（クラス）を選択して頭髪色領域を抽出する。このときのクラス分割手法としては、前述のようにK-平均アルゴリズムなどの一般に知られているクラスタリングのアルゴリズムを使用して差し支えない。さらに、分割クラス数は実験的に求めた最適な値を使用する。

【0031】そして、あらかじめ求めておいた頭髪領域の全画面に対する面積比率により、上記で求まった頭髪候補領域のうちで、前述した人物顔領域抽出部3で抽出した顔領域よりも画面上方に位置するもので、顔と面積比率があらかじめ実験的に求めて登録してある頭髪のそれと最も近いものを選択することで、頭髪領域を抽出する。

【0032】ここでの頭髪領域の面積の計算は、前述の

$$\theta = 1 / 2 \cdot \tan^{-1} (2 \cdot m_{11} / (m_{02} - m_{20})) \dots (3)$$

ここで、 $m_{11}$ ,  $m_{02}$ ,  $m_{20}$ は、それぞれ $m_{11} = \iint xy f(x, y) dx dy$ ,  $m_{20} = \iint x^2 f(x, y) dx dy$ ,  $m_{02} = \iint y^2 f(x, y) dx dy$ である。

【0037】次に、第1の顔特徴点抽出部6の処理について、図6に示すフローチャートを参照して説明する。前述したように、第1の顔特徴点としては口を選択し、これを抽出する。

【0038】一般に、YIQ表色系のQ成分画像においては、口の領域の濃淡値は特定値付近に集中することが知られている。この性質を利用して、入力画像のうち、前述の人物顔領域抽出部3の処理で抽出された顔領域画像の内部について、Q成分画像の濃度ヒストグラムを計測し、クラス分けする。さらに、このクラスのうち、口領域を端的に示すものをあらかじめ教示しておき、その領域を抽出する。ここでは、分割クラス数は実験的に求めたものを使用し、このうち口領域として抽出するクラ

$$E_m = -(W_s \cdot S + W_p \cdot P + W_{pp} \cdot PP) \dots (4)$$

である。ただし、Sは候補領域の面積、Pは候補領域の高さ方向の位置（大なる程、画面下に位置する）、PP

人物顔領域抽出部3と同様の手法を使用して差し支えない。上述の処理によって抽出された頭髪領域は、頭髪部分の濃淡値を「1」、それ以外の部分の濃淡値を「0」とした2値デジタル画像（頭髪領域画像）として出力される。

【0033】次に、人物頭部領域抽出部5の処理について、図4に示すフローチャートを参照して説明する。上述した人物顔領域抽出部3および人物頭髪領域抽出部4の各処理によって、画面内的人物の顔領域画像、および、頭髪領域画像の領域を切出すことができ、両領域は前述のデジタル画像として表現されるため、本人物頭部領域抽出部5の処理では、まず、これらを図5(a)に示す画像の論理和演算によって合成し、頭部領域画像を作成する。

【0034】次に、合成された領域について、以下の式(2)にしたがって、その重心座標を求める。すなわち、人物頭部領域の2値デジタル画像を $f(x, y)$ とすると、重心座標 $(x_g, y_g)$ は、 $x$ ,  $y$ 軸方向への1次モーメントであり、

$$x_g = \iint x f(x, y) dx dy$$

$$y_g = \iint y f(x, y) dx dy \dots (2)$$

で示される。

【0035】また、上記合成領域の2次のモーメントから、慣性主軸方向を求め、これを投影平面内での頭部傾き角度 $\theta$ とする（図5b参照）。すなわち、 $x$ ,  $y$ 軸方向の0, 1, 2次の慣性モーメントを $m_{ij}$  ( $i, j = 0, 1, 2$ ) とすると、慣性主軸方向（頭部傾き角度） $\theta$ は以下の式(3)にしたがって算出する。

【0036】

$$\theta = 1 / 2 \cdot \tan^{-1} (2 \cdot m_{11} / (m_{02} - m_{20})) \dots (3)$$

スは、濃淡値が最大になるものとした。

【0039】したがって、前述の人物顔領域抽出部3と同様に、まず、入力画像に対してRG B→YIQ変換を行なう。次に、上述の処理にて抽出されたQ成分画像のみの濃度ヒストグラムをとり、これを濃度軸上で領域分割し、濃度が最大の領域（クラス）を選択して口候補領域を抽出する。ここでも、クラスタリング手法は、前述同様、既存のものを使用して差し支えない。

【0040】ここまで処理によって、口領域の候補領域が抽出されるわけであるが、口以外の領域についても抽出されてしまう可能性があるため、各候補領域の外接矩形の内部に対して以下の評価関数値を最小化するものを選択して、これらを限定し、口領域周囲の外接矩形のみを抽出し、その重心座標を求める。すなわち、評価関数 $E_m$ は、

$$E_m = -(W_s \cdot S + W_p \cdot P + W_{pp} \cdot PP) \dots (4)$$

は候補領域内の最大幅長、 $W_s$ ,  $W_p$ ,  $W_{pp}$ は重み定数である。

【0041】次に、第2の顔特徴点抽出部7の処理について、図7に示すフローチャートを参照して説明する。前述したように、第2の顔特徴点としては目を選択し、これを抽出する。

【0042】一般に、カラー画像を下記数1の色相(H)、彩度(S)(文献[7] 杉山、田辺：動画像を用いた実時間瞬目抽出法、画像ラボ、1993.9、PP.45-48参照)で表現したときのS成分においては、目の瞳領域の濃淡値は特定値付近に集中することが知られている。

【0043】

【数1】

$$H = \tan^{-1} ((r - 85) / (g - 85))$$

$$S = \sqrt{(r - 85)^2 + (g - 85)^2}$$

【0044】ただし、R、G、Bは0~255とし、 $r = 255R / (R+G+B)$ 、 $g = 255G / (R+G+B)$ とする。この性質を利用して、入力画像のうち、前述の人物顔領域抽出部3の処理で抽出された顔領域画像の内部について、S成分画像の濃度ヒストグラムを計測し、クラス分けする。さらに、このクラスのうち、瞳領域を端的に示すものをあらかじめ教示しておき、その領域を抽出する。

【0045】しかし、この変換においても、たとえば、髪の毛が目の部分にかかっている場合は、瞳、髪のS成分画像での濃度が接近しているため、分離が困難となる。そのため、ここでは、S成分画像に対して、評価関数値の最小化による入力顔画像適応形のテンプレートマッチング手法を用いて、目領域のみの抽出精度を高めるようとする。

【0046】マッチングの特徴量は、以下のような目の領域に関する周囲領域、位置、形状の3つの性質を使用する。

$$E_e = (W_n \cdot N + W_d \cdot D + W_f \cdot F) \dots (6)$$

ただし、Nは注目ブロックの周囲の瞳濃度クラスに属する画素数、Dは顔の対象軸からの距離(対象軸との傾きで代用)、Fは平均的な人物における瞳領域と目領域との面積比からの誤差、Wn、Wd、Wfは重み定数である。

【0053】次に、方向算出部8の処理について説明する。方向算出部8は、顔全体の位置に対する顔内部の特徴点との位置関係によって顔の方向を推定する。これには様々な既存手法が考えられるが、本実施の形態では、従来の技術で説明した頭全体の重心と顔内部の特徴点として高輝度変化領域を抽出し、それらの重心同士のずれを比較する方式(文献[3] 参照)を応用している。

【0054】すなわち、頭全体としては、前述の人物頭部領域抽出部5で抽出された頭部領域を用い、顔内部の高輝度変化領域の重心の代わりに、前述の第1の顔特徴

(a) 周囲領域：画像中の目領域の周囲の画素の濃淡値は瞳領域のそれとは異なる。

【0047】(b) 位置：目領域は顔の対象軸からあまり離れていない。

(c) 形状：瞳領域と目の全体領域それぞれとの顔領域に対する幅の比率、および、瞳、目全体それぞれの面積比率はほぼ一定である。これは、複数の人物をサンプルとするカラー顔画像の計測により求める。

【0048】したがって、これに基づくマッチングアルゴリズムは次に示すようになる。まず、前述の人物顔領域抽出部3の処理で抽出された顔領域の高さHfおよび幅Wfから、目の領域の高さHeおよび幅Weを下記式(5)によって推定する。

【0049】

$$He = Hf \cdot Rh$$

$$We = Wf \cdot Rw \dots (5)$$

ここで、RhおよびRwは、それぞれ顔高さと目高さとの比率および顔幅と目幅との比率であり、複数の人物サンプルにおける顔高さおよび幅と目のそれらとの比率の平均的な値を用いる。

【0050】次に、顔領域内部において、前記数1におけるS成分画像を求め、口の抽出と同様、濃淡ヒストグラムを分割し、その内の特定クラスを抽出する。このときの分割クラス数は実験的に求め、特定クラスは瞳の濃度に相当するもので、分割クラスのうち濃度の最小なものとした。

【0051】次に、この濃度クラスのみ抽出された画像のうち、顔の上方に位置し、顔幅に相当する領域を、式(5)の目領域の大きさに相当するブロックに分割し、各ブロック内部での下記式(6)に示す評価関数値Eeが最小なるものを選択し、その評価値最小ブロックについて重心座標を求める。

【0052】

$$E_e = (W_n \cdot N + W_d \cdot D + W_f \cdot F) \dots (6)$$

点抽出部6、および、第2の顔特徴点抽出部7で抽出された各特徴点の位置の重心を使用している。

【0055】次に、人物操作状態算出部9の処理について、図8に示すフローチャートを参照して説明する。人物操作状態算出部9は、上述の顔方向算出部8によって算出された顔方向を一定の離散時間で求め、これらの時間的な変化パターンや統計量を算出し、あらかじめ複数の被験者を対象として求められ、被験者の状態(注目している、操作に迷っている、操作を中断している等)を示す標準的な参照データと比較することにより、操作人物の状態を推定する。

【0056】すなわち、本実施の形態においては、顔方向の時間的な統計量としては、顔方向の時間平均θaと時間分散θv、および、顔方向が指定操作範囲(たとえば操作パネル上)外を示した時間θxを求める。そし

て、これらを、あらかじめ求めておいた、それぞれの閾値  $T_a$ 、 $T_v$  および  $T_x$  と比較して、以下の 3 種類の状

$$\begin{aligned} \theta_a > T_a &\text{かつ } \theta_v < T_v \\ \theta_a < T_a &\text{かつ } \theta_v > T_v \\ \theta_x > T_x \end{aligned}$$

ここで、出力部 10 へは、上記  $\theta_a$ 、 $\theta_v$ 、および、上記 3 状態をステータス信号として出力する。

【0058】次に、出力部 10 の処理について説明する。出力部 10 は、全体処理でも説明した通り、算出された顔方向のベクトルに注目度に応じた重みを伴わせたデータと、操作状態のデータ（注目している、操作に迷っている、操作を中断している）として出力される。ここで、注目度  $N$  としては、上述の  $\theta_a$ 、 $\theta_v$  を用いて、 $N = W_a \cdot \theta_a + W_v \cdot \theta_v \dots (8)$  となる。

【0059】ここに、 $W_a$ 、 $W_v$  は、表示階調を注目度に比例するようにスケーリングするための定数である。また、操作状態の出力としては、人物操作状態算出部 9 の処理で示した 3 状態のステータス信号をそのまま使用する。

【0060】次に、人物注目方向によるグラフィカルユーザインターフェース（GUI）について説明する。図 9 は、図 1 における符号 1～10 までを 1 つの処理装置、すなわち、人物注目方向検出装置 11 と考え、出力部 10 からのデータを、グラフィカルユーザインターフェース装置としての、表示装置 12 を備えた汎用のパーソナルコンピュータ（汎用パソコン）13 に与えることによって、操作人物の顔の方向と注目度とに応じた、表示装置 12 の画面内のアイコンクリック、および、操作メッセージ起動の例を説明する図である。

【0061】人物注目方向検出装置 11 からパーソナルコンピュータ 13 に送られたステータス信号（人物操作状態算出部 9 の処理で述べたもの）のそれぞれの状態に応じて、パーソナルコンピュータ 13 の表示装置 12 では、以下に示す (a)～(c) のようなアクションをおこし、ユーザとの会話を行なう。このときの画像入力部 1 のビデオカメラ 1a と表示装置 12 の画面との位置関係を図 10 に示す。

【0062】(a) ステータス信号＝「注目している」の場合

前述の顔重心位置 14 を原点とする座標系を考え、そこから表示装置 12 の画面に向かう前述の方向ベクトルと画面との交点 ( $x_i$ ,  $y_i$ ,  $z$ ) を求め、その位置に最も近く、現在のユーザ（操作人物）15 のアクションとして可能性のあるアイコン位置をユーザ 15 によるピック位置とし、対応するアクションを起す。

【0063】たとえば、ユーザ 15 のイエス／ノーの選択時には、ユーザ 15 のピックの探索は、これら 2 つのアイコンが、ユーザ 15 に選択されたものとしての可能性がある。なお、Z は、あらかじめ設定しておいたユー

態を決定する。

#### 【0057】

→注目している。

→操作に迷っている。

→操作を中断している。 …… (7)

ザ 15 の顔と表示装置 12 の画面との距離である。

【0064】(b) ステータス信号＝「操作に迷っている」の場合

ユーザ 15 に対して、たとえば、表示と音声によって「お困りですか？」というメッセージを出し、さらに、たとえば、「上記の質問に答えるためにイエス／ノーの位置を注視してください。」といった、操作のより詳しいメッセージを表示する。

【0065】(c) ステータス信号＝「操作を中断している」の場合

中断している状態が、非常に長い一定回数だけ続いた場合には、表示装置 12 の画面の焼き付きを保護するためにスクリーンセイバープログラムを起動する。このプログラムについては、どのようなものを使用してもよい。

#### 【0066】

【発明の効果】以上詳述したように本発明の人物注目方向検出方式によれば、入力された操作人物の顔領域を含む画像から操作対象人物の顔領域および顔内部の特徴点位置を特定し、それらを用いて顔の方向とその時間的な統計を検出することによって操作人物の注目方向と感情を含む操作状態を推定することにより、操作人物の状態を自動的に把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示などを行なうことが可能となる。

【0067】すなわち、たとえば、現金自動入出金装置や券売機などの会社システム機器、あるいは、パーソナルコンピュータやワードプロセッサなどのオフィスオートメーション機器などにおいて、一般ユーザが操作方法を理解できないための問題があるが、本発明方式を用いることにより、操作人物の状態を機器側が把握し、操作人物の状態に応じてユーザフレンドリな操作方法の表示（ヘルプメニュー）などを行なうことが可能となる。

【0068】また、現在、パーソナルコンピュータやワークステーションなどのグラフィカルユーザインターフェース装置は、グラフィカルユーザインターフェース（GUI）を備えていて、マウス操作を主体とした操作ができるが、本発明方式を用いることにより、注目領域による操作ボタンの処理や、注目領域の表示を変更（拡大、表示色の変更など）することによる、高度なグラフィカルユーザインターフェースが構築可能となる。

#### 【図面の簡単な説明】

【図 1】本発明の実施の形態に係る人物注目方向検出方式が適用される装置の構成を示すブロック図。

【図 2】人物顔領域抽出部の処理を説明するフローチャート。

【図3】人物頭髪領域抽出部の処理を説明するフローチャート。

【図4】人物頭部領域抽出部の処理を説明するフローチャート。

【図5】人物頭部領域画像の合成および人物頭部領域画像の投影面内での傾き角度を説明する図。

【図6】第1の顔特徴点抽出部の処理を説明するフローチャート。

【図7】第2の顔特徴点抽出部の処理を説明するフローチャート。

【図8】人物操作状態算出部の処理を説明するフローチャート。

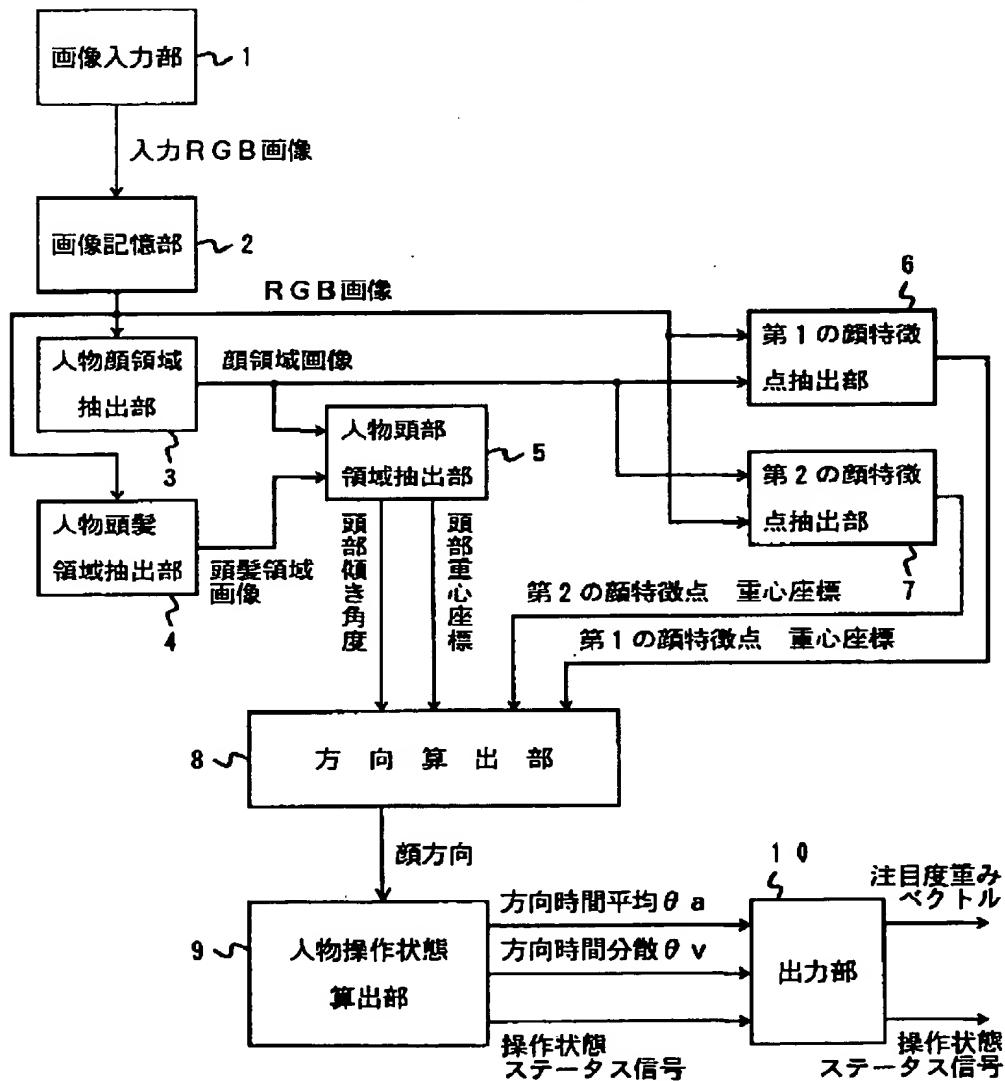
【図9】人物注目方向によるグラフィカルユーザインターフェースについて説明するブロック図。

【図10】画像入力部のビデオカメラと表示装置の画面との位置関係を示す概略図。

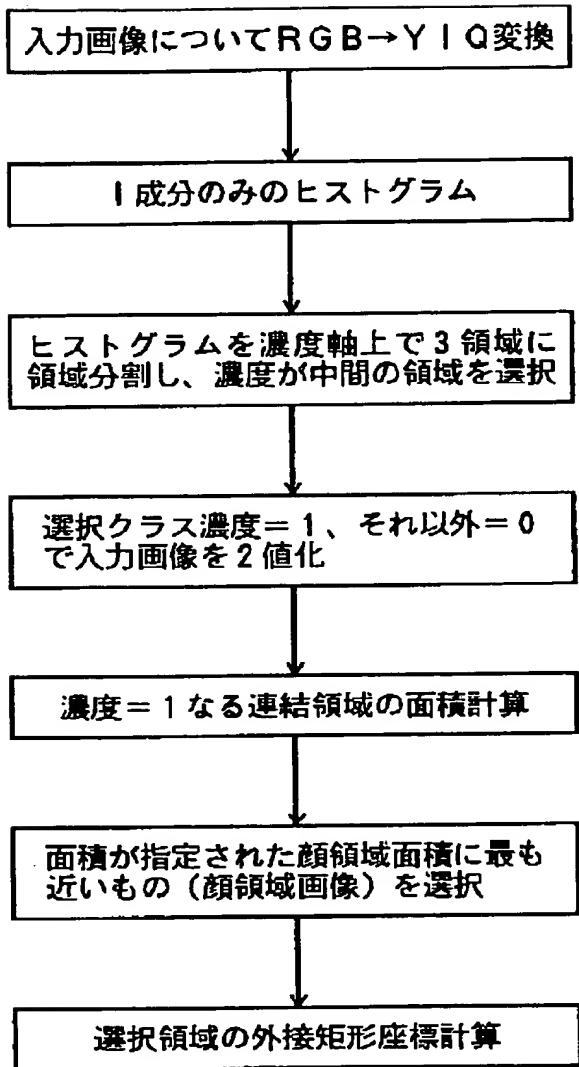
#### 【符号の説明】

1……画像入力部、2……画像記憶部、3……人物顔領域抽出部、4……人物頭髪領域抽出部、5……人物頭部領域抽出部、6……第1の顔特徴点抽出部、7……第2の顔特徴点抽出部、8……方向算出部、9……人物操作状態算出部、10……出力部、11……人物注目方向検出装置、12……表示装置、13……パーソナルコンピュータ。

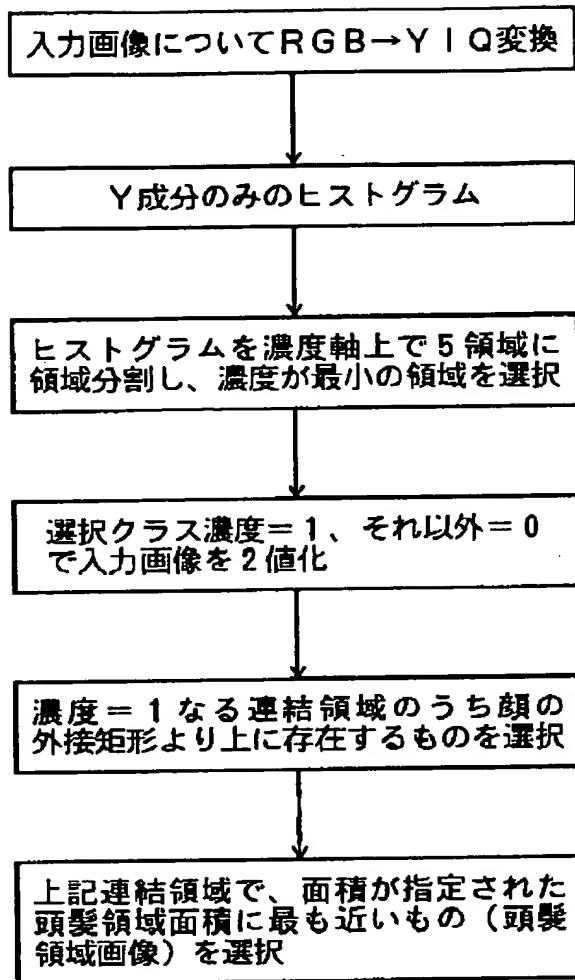
【図1】



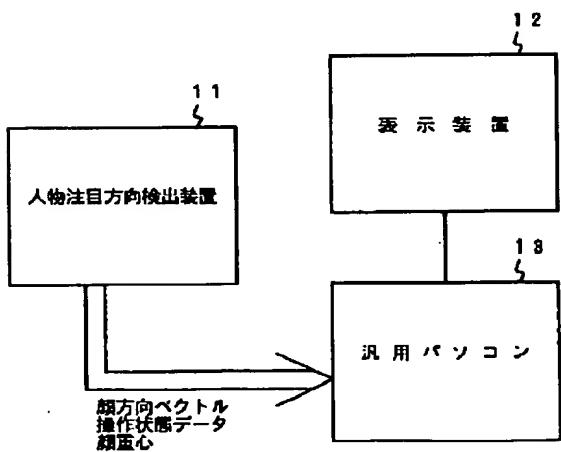
【図2】



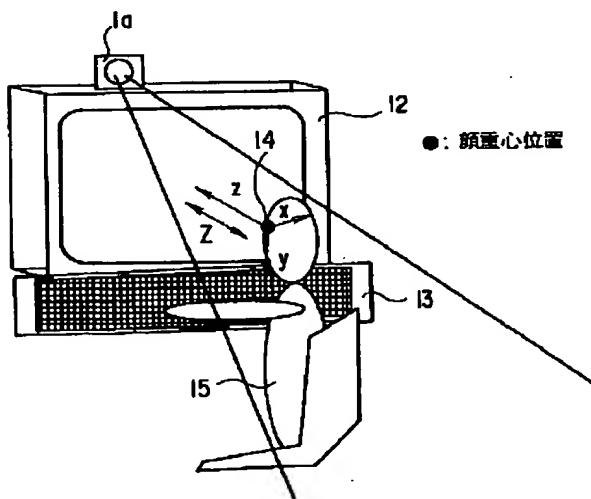
〔図3〕



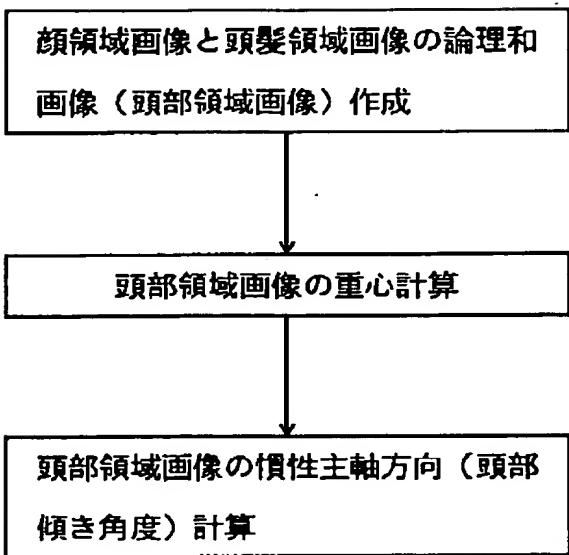
〔図9〕



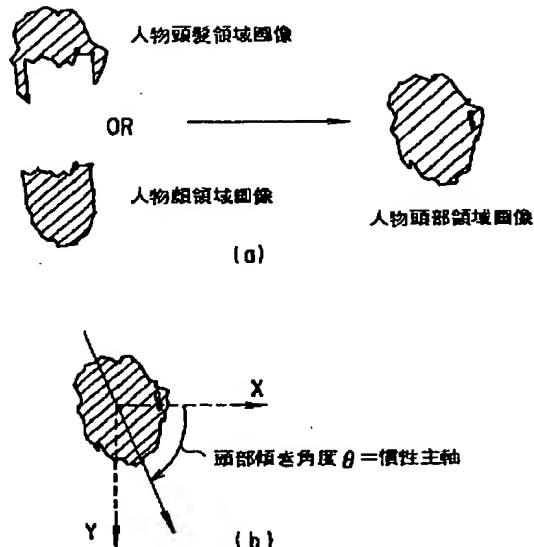
【四】



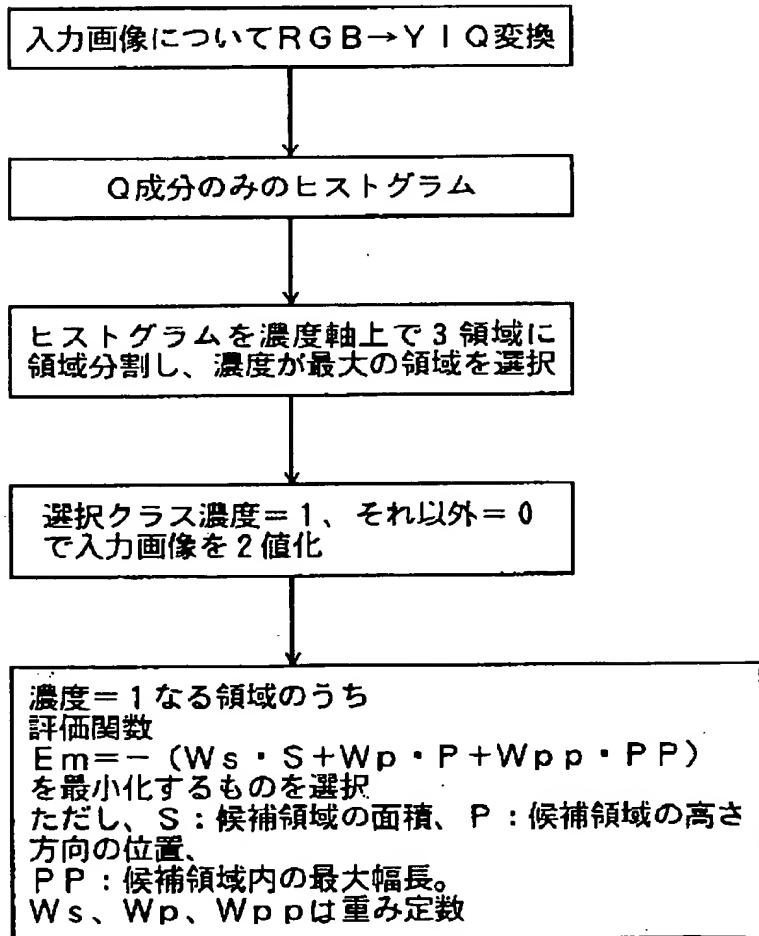
【図4】



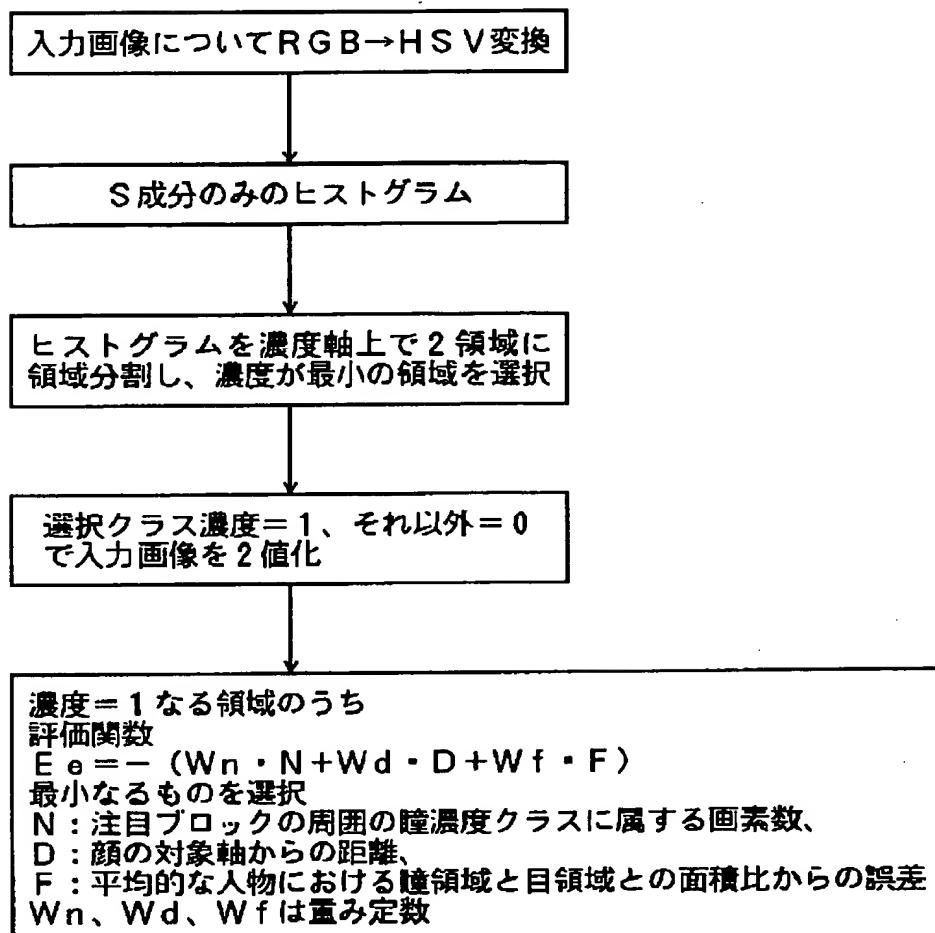
【図5】



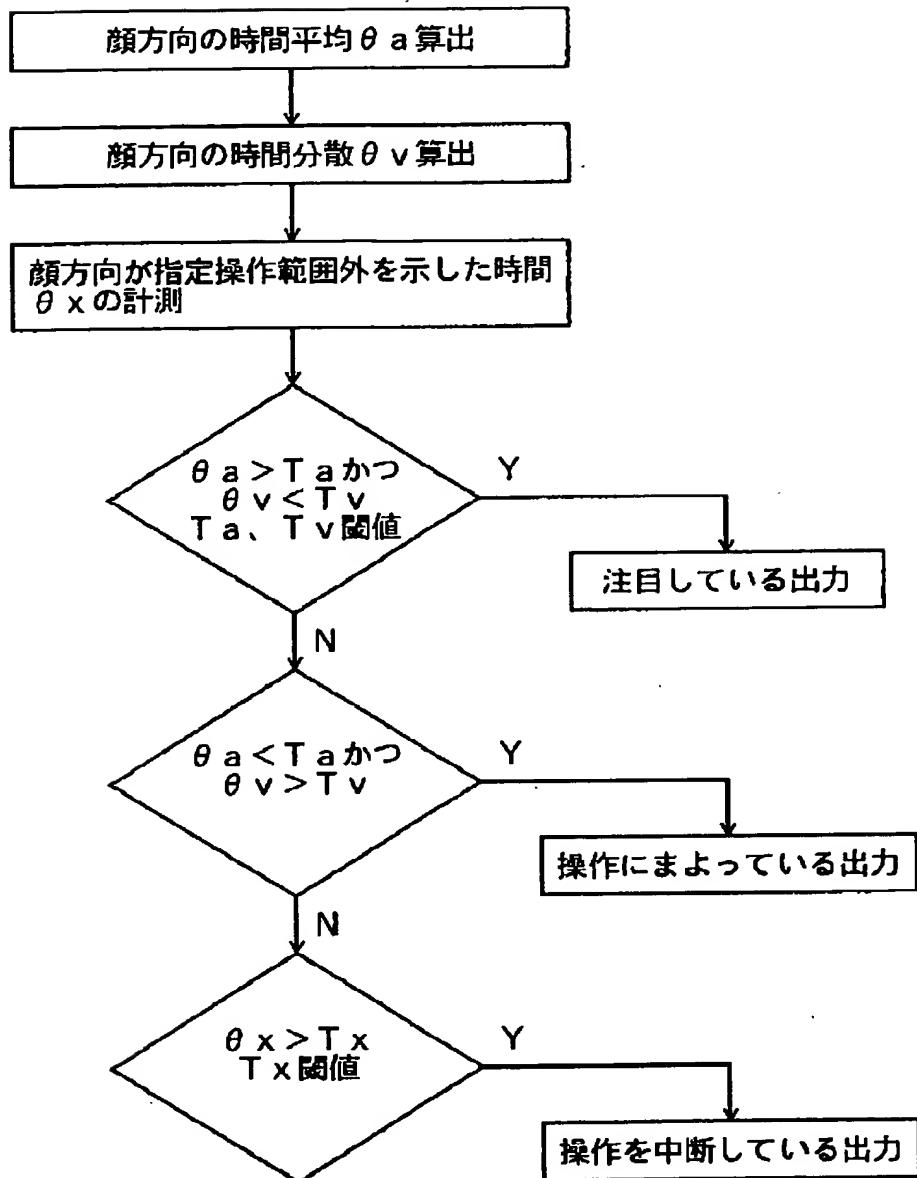
【図6】



【図7】



【図8】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**